AD_____

Award Number: DAMD17-00-1-0245

TITLE: Robust Detection of Masses in Digitized Mammograms

PRINCIPAL INVESTIGATOR: Lihua Li, Ph.D.

CONTRACTING ORGANIZATION: University of South Florida
Tampa, FL 33620

REPORT DATE: May 2005

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 2005	Annual Summary (15 May 00 - 14 May 05)

4. TITLE AND SUBTITLE
Robust Detection of Masses in Digitized Mammograms

5. FUNDING NUMBERS
DAMD17-00-1-0245

6. AUTHOR(S)
Lihua Li, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of South Florida
Tampa, FL 33620

E-Mail: lilh@moffitt.usf.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

This project is to develop a robust computer aided diagnosis (CAD) system for mass detection with high sensitivity and specificity in digitized mammograms. The research in the year of non-cost extension is on data collection, evaluation of clinical significance of CAD system, and the final documentation and manuscript reports. By reviewing more than 1553 cases, a total of 100 consecutive cancer cases were collected. A ground truth file was generated by an experienced radiologist. To evaluate the clinical significance of the CAD system in early cancer detection, a testing of the CAD system on independent consecutive cancer cases was performed, in which the detection FROC curves and months of early detection at different false-positive rates were obtained.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast cancer, mass, detection, CAD, mammography, robust, adaptive, segmentation, classification	9
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

INTRODUCTION

This project is to develop a robust computer aided diagnosis (CAD) system for mass detection with high sensitivity and specificity in digitized mammograms. As listed in the Statement of Work, the research scope in the fourth year of project is to evaluate the clinical significance of CAD system and finish the final documentation and manuscript reports. However, due to the unexpected difficulty and the huge working load in serial data collection, the research planned for fourth year was delayed to the fifth year with no-cost extension. This study was taken by testing how well the CAD algorithm performs in early detection of masses with a consecutive set of mammograms.

BODY

Objective 1: ***to generate a database containing 100 consecutive cases for evaluation of clinical significance of CAD system.***

Accomplishments:

1. Data Collection Criteria and Procedure

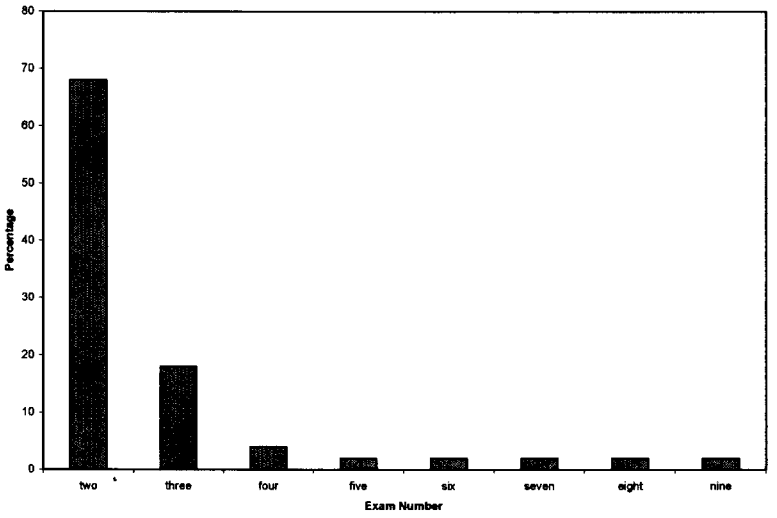
In order to test the clinical significance of CAD system in terms of early detection, the collection of a serial data set of mammograms was continued based on fourth year's work. As we used before, the criteria for inclusion of mammogram in this study are as follows:

1. Mass must be visible on mammogram
2. Mass must be proven by biopsy to be malignant
3. Mass must be seen in retrospect on a prior mammogram when reviewed by a radiologist

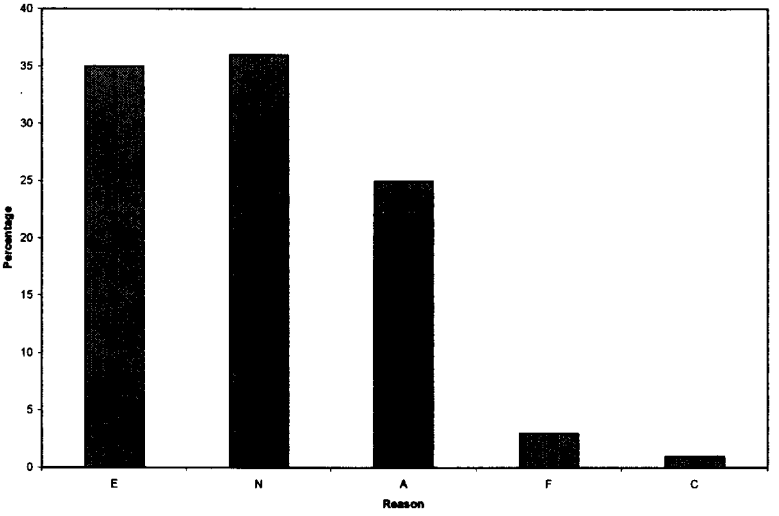
A hundred (100) missed cancer cases were selected by reviewing 1553 consecutive mammogram cases, from which three datasets were generated including mammograms with missed cancer, mammograms with screening-detected cancer and normal mammograms. The radiologist indicated the location and outlined the contour of the lesion on both exams and the BIRADS descriptors. Ground truth files (hard copy) were generated based on the radiologist's outlines. The films were then digitized on a Kodak (LUMISYS) LS85 digitizer at a resolution of 50μm and 12 bits in grey scale.

2. Characteristic analysis of the database

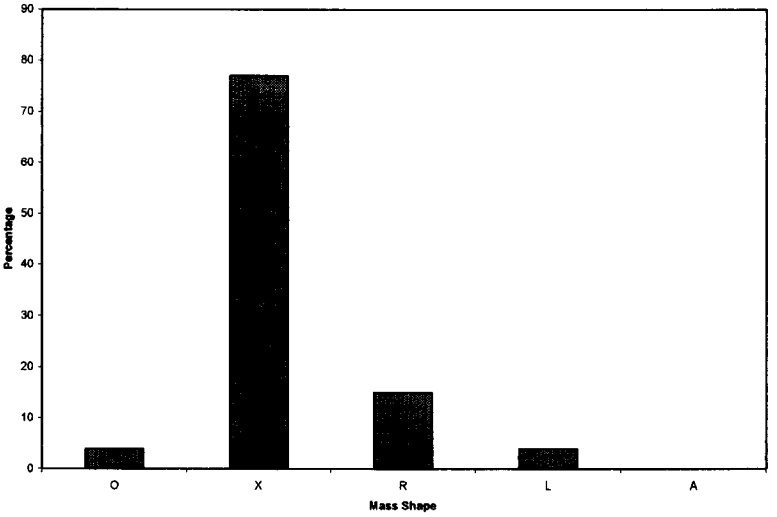
The characteristics of database was updated here by following descriptions: (a) Case distribution in terms of exam number, (b) Case distribution in terms of missing reason, (c) Case distribution in terms of mass shape, (d) Case distribution in terms of mass margin, (e) Case distribution in terms of mass density. The histograms are shown in Figure 1.


80
70
60
50
40
30
20
10
0
Percentage
two
three
four
five
six
seven
eight
nine
Exam Number

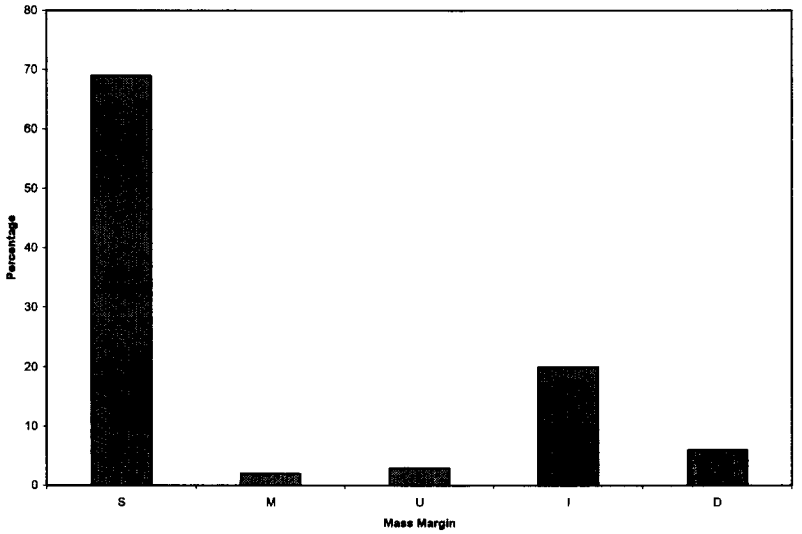
(a)


40
35
30
25
20
15
10
5
0
Percentage
E
N
A
F
C
Reason

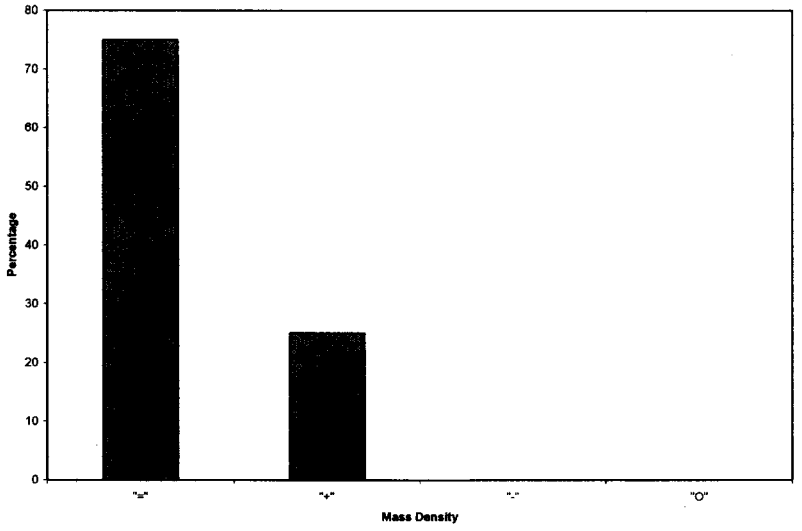
(b)


90
80
70
60
50
40
30
20
10
0
Percentage
O
X
R
L
A
Mass Shape

(c)



(d)



(e)

Figure 1. Case distribution in terms of (a) exam numbers, (b) missing reason, (b) mass shape, (d) mass margin, (e) mass density.

Objective 2: ***evaluation of clinical significance of CAD system in terms of early detection***

To evaluate the clinical significance of the CAD system in early cancer detection, a testing of the CAD system on consecutive cancer cases was performed. Here the consecutive cancer cases are completely independent of the database used for CAD system training and design. The characteristics of the database are described above. Figure 2 is the distribution of interval between the mammograms taken at detection and missed stages.

The FROC curves of detection on mammograms at missed and detected stages are shown in Figure 3. Each curve was generated by five pairs of sensitivity/false-positive detections. It is observed that there is a big difference in detection performance between detection and missed stages, which to some extent shows the difficulty in cancer screening.

The early detection performance is evaluated using the average number of months of detection of missed cancers by CAD before radiologist's detection. Specifically, if a cancer at missed stage could not be detected by CAD, the month of early detection is counted as 0, otherwise equals to

the interval between missed and detected stages for that case. Figure 4 shows the number of months of early detection at different false-positive rates.

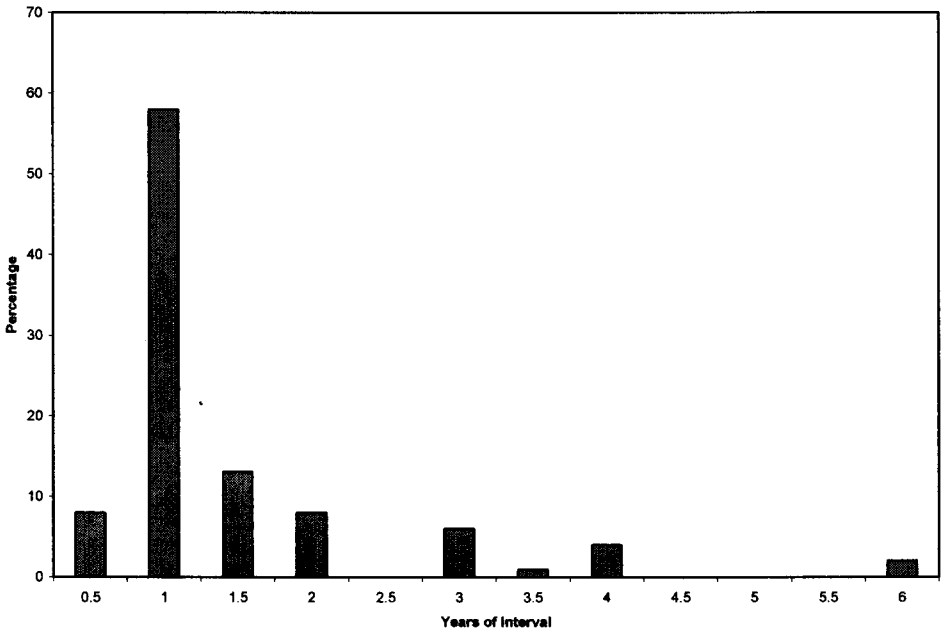


Figure 2. Distribution of interval between the mammograms taken at detection and missed stages

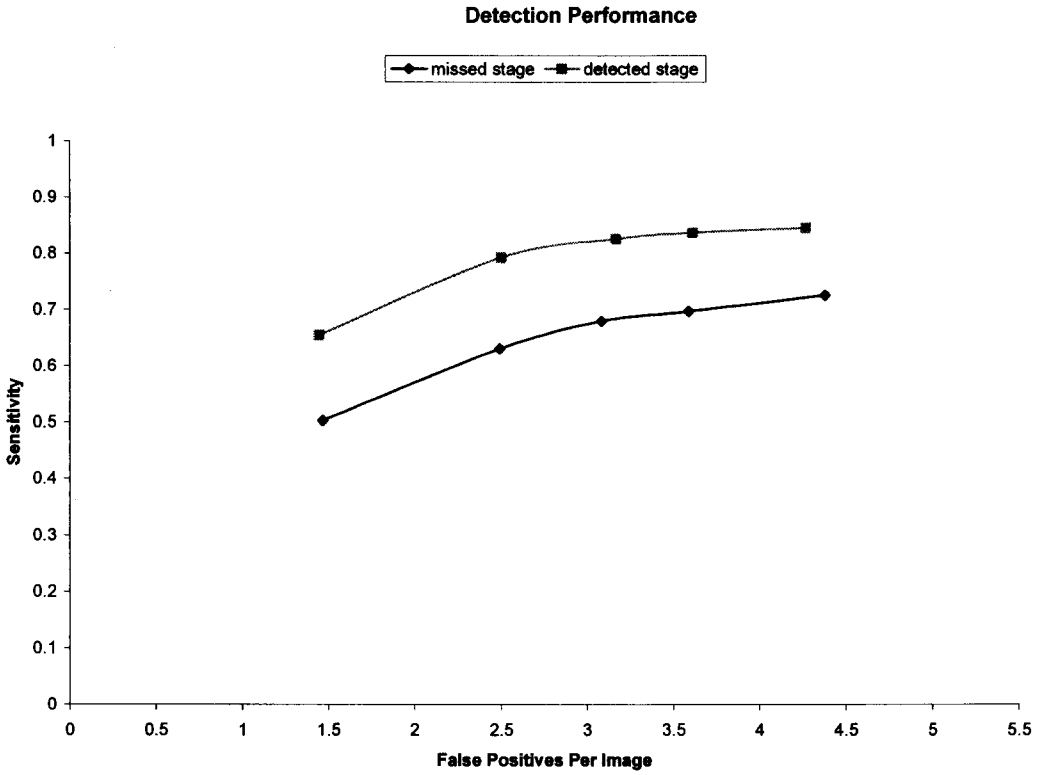


Figure 3. FROC curves of detection on mammograms at missed and detected stages.

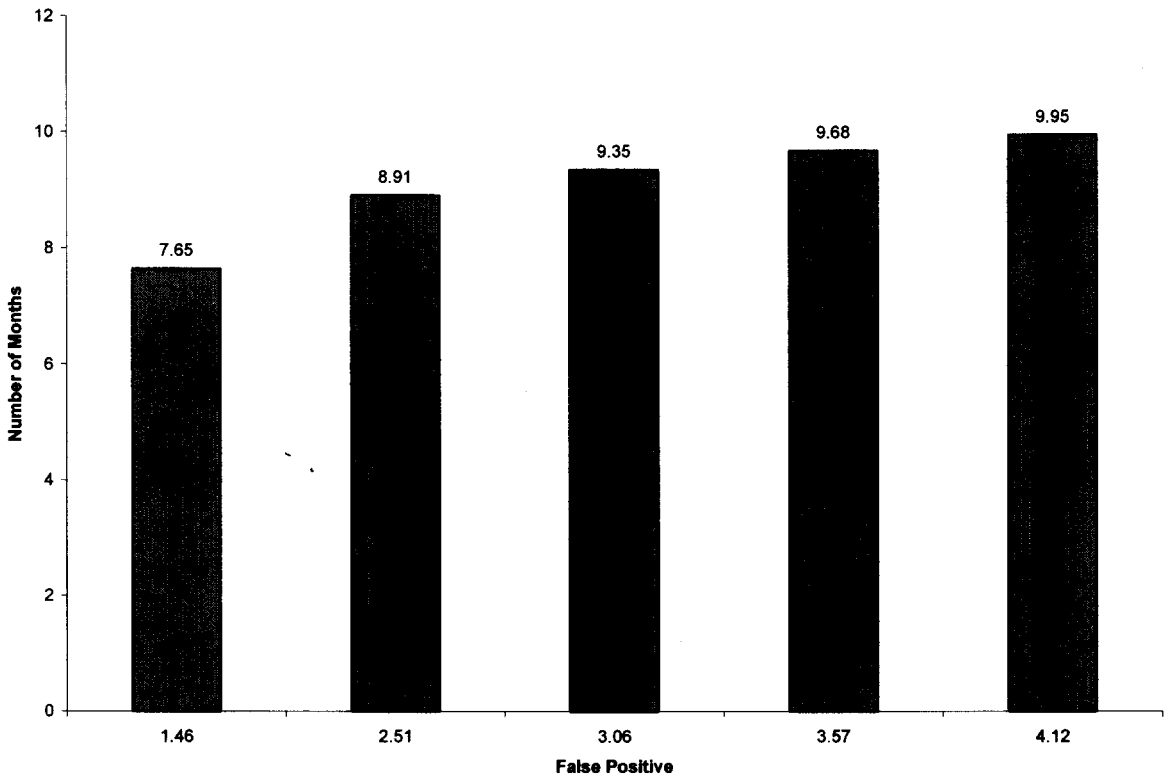


Figure 4. The number of months of early detection by CAD at different false-positive rates.

KEY RESEARCH ACCOMPLISHMENTS

1. A database of mammogram was generated containing 100 cases of serial mammograms, which were selected by reviewing more than 1553 cases.
2. An analysis of collected database was taken in terms of consecutive exam number, mass shape, mass margin, mass density, and the interval between the mammograms taken at detection and missed stages.
3. To evaluate the clinical significance of the CAD system in early cancer detection, a testing of the CAD system on independent consecutive cancer cases was performed. The number of months of early detection at different false-positive rates was obtained.

REPORTABLE OUTCOMES

1. Presentation and/or proceedings paper

(a) Lihua Li, Zuobao Wu, Florence George, Zhao Chen, Angela Salem, Maria Kallergi, Claudia G. Berman, "Breast Tissue Density and CAD Cancer Detection in Digital Mammography," Proceedings of 27th IEEE Conference on EMBS, 2005.

CONCLUSIONS

This project is to develop a robust computer aided diagnosis (CAD) system for mass detection with high sensitivity and specificity in digitized mammograms. The research in this fifth year of

non-cost extension is to finish the research work that could not have been done in fourth year due to the difficulty in data collection. The research progresses include collection of 100 consecutive cases, evaluation of clinical significance of CAD system, and the final documentation and manuscript reports. By reviewing more than 1553 cases, a total of 100 serial cancer cases were collected. A ground truth file was generated by an experienced radiologist. To evaluate the clinical significance of the CAD system in early cancer detection, a testing of the CAD system on independent consecutive cancer cases was performed. It is observed that (1) on average the cancer could be detected 7.65 months earlier before radiologist detection by using CAD at a false positive of 1.46 per image; (2) the sensitivity of CAD on missed cancer detection is still relatively low. How to improve the detection of missed cancer is a challenging task for CAD system design, which is a topic this PI and his colleagues are working on.